Regional Operational Plan CF.5J.14.01

Operational Plan: Cook Inlet Coho Salmon Genetic Baseline Study

by

Andrew W. Barclay

Nicholas A. DeCovich

and

Christopher Habicht

March 2014

Alaska Department of Fish and Game **Divisions of Sport Fish and Commercial Fisheries**

Symbols and Abbreviations

The following symbols and abbreviations, and others approved for the Système International d'Unités (SI), are used without definition in the following reports by the Divisions of Sport Fish and of Commercial Fisheries: Fishery Manuscripts, Fishery Data Series Reports, Fishery Management Reports, and Special Publications. All others, including deviations from definitions listed below, are noted in the text at first mention, as well as in the titles or footnotes of tables, and in figure or figure captions.

Weights and measures (metric)

centimeter	cm
deciliter	dL
gram	g
hectare	ha
kilogram	kg
kilometer	km
liter	L
meter	m
milliliter	mL
millimeter	mm

Weights and measures (English)

cubic feet per second	ft^3/s
foot	ft
gallon	gal
inch	in
mile	mi
nautical mile	nmi
ounce	oz
pound	lb
quart	qt
yard	yd

Time and temperature

day	d
degrees Celsius	°C
degrees Fahrenheit	°F
degrees kelvin	K
hour	h
minute	min
second	s

Physics and chemistry

all atomic symbols	
alternating current	AC
ampere	A
calorie	cal
direct current	DC
hertz	Hz
horsepower	hp
hydrogen ion activity (negative log of)	pH
parts per million	ppm
parts per thousand	ppt, ‰
volts	V
watts	W

General

Alaska Administrative Code	AAC
all commonly accepted abbreviations	e.g., Mr., Mrs., AM, PM, etc.
all commonly accepted professional titles	e.g., Dr., Ph.D., R.N., etc.
at	@
compass directions:	
east	E
north	N
south	S
west	W
copyright	©
corporate suffixes:	
Company	Co.
Corporation	Corp.
Incorporated	Inc.
Limited	Ltd.
District of Columbia	D.C.
et alii (and others)	et al.
et cetera (and so forth)	etc.
exempli gratia (for example)	e.g.
Federal Information Code	FIC
id est (that is)	i.e.
latitude or longitude	lat. or long.
monetary symbols (U.S.)	$, ¢
months (tables and figures): first three letters	Jan,...,Dec
registered trademark	®
trademark	™
United States (adjective)	U.S.
United States of America (noun)	USA
U.S.C.	United States Code
U.S. state	use two-letter abbreviations (e.g., AK, WA)

Mathematics, statistics

all standard mathematical signs, symbols and abbreviations	
alternate hypothesis	H_A
base of natural logarithm	*e*
catch per unit effort	CPUE
coefficient of variation	CV
common test statistics	(F, t, χ^2, etc.)
confidence interval	CI
correlation coefficient (multiple)	R
correlation coefficient (simple)	r
covariance	cov
degree (angular)	°
degrees of freedom	df
expected value	*E*
greater than	>
greater than or equal to	≥
harvest per unit effort	HPUE
less than	<
less than or equal to	≤
logarithm (natural)	ln
logarithm (base 10)	log
logarithm (specify base)	$\log_2$, etc.
minute (angular)	'
not significant	NS
null hypothesis	H_O
percent	%
probability	P
probability of a type I error (rejection of the null hypothesis when true)	α
probability of a type II error (acceptance of the null hypothesis when false)	β
second (angular)	"
standard deviation	SD
standard error	SE
variance	
population	Var
sample	var

REGIONAL OPERATIONAL PLAN CF.5J.14.01

COOK INLET COHO SALMON GENETIC BASELINE STUDY

by

Andrew W. Barclay

Nicholas A. DeCovich

and

Christopher Habicht

Alaska Department of Fish and Game, Division of Commercial Fisheries, Anchorage

Alaska Department of Fish and Game
Division of Commercial Fisheries

March 2014

The Regional Operational Plan Series was established in 2012 to archive and provide public access to operational plans for fisheries projects of the Divisions of Commercial Fisheries and Sport Fish, as per joint-divisional Operational Planning Policy. Documents in this series are planning documents that may contain raw data, preliminary data analyses and results, and describe operational aspects of fisheries projects that may not actually be implemented. All documents in this series are subject to a technical review process and receive varying degrees of regional, divisional, and biometric approval, but do not generally receive editorial review. Results from the implementation of the operational plan described in this series may be subsequently finalized and published in a different department reporting series or in the formal literature. Please contact the author if you have any questions regarding the information provided in this plan. Regional Operational Plans are available on the Internet at: http://www.adfg.alaska.gov/sf/publications/

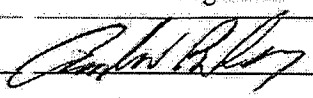
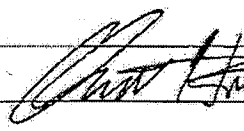
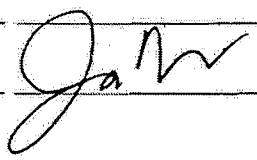
Andrew W. Barclay
Nicholas A. DeCovich
and
Christopher Habicht
Alaska Department of Fish and Game, Division of Commercial Fisheries,
333 Raspberry Road, Anchorage, Alaska, 99518-1599

This document should be cited as:
Barclay, A. W., N. A. DeCovich and C. Habicht . 2014. Cook Inlet coho salmon genetic baseline study. Alaska Department of Fish and Game, Regional Operational Plan CF.5J.14.01, Anchorage.

SIGNATURE PAGE

Project Title:	Cook Inlet Coho Salmon Genetic Baseline Study
Project leader(s):	Andrew W. Barclay Fishery Biologist III, Nicholas A. DeCovich Fishery Biologist III
Division, Region, and Area	Commercial Fisheries, Region V, Headquarters, Juneau
Project Nomenclature:	
Period Covered	
Field Dates:	
Plan Type:	Category II

Approval

Title	Name	Signature	Date
Project leader	Andrew Barclay		3/3/14
Project leader	Nicholas Decovich		
Biometrician	Jim Jasper		
Research Coordinator	Chris Habicht		3/3/14
Region II research biologist	Mark Willette		
Sport Fish Regional Research Supervisor	Jack Erickson		3/3/14

SIGNATURE PAGE

Project Title:	Cook Inlet Coho Salmon Genetic Baseline Study
Project leader(s):	Andrew W. Barclay Fishery Biologist III, Nicholas A. DeCovich Fishery Biologist III
Division, Region, and Area	Commercial Fisheries, Region V, Headquarters, Juneau
Project Nomenclature:	
Period Covered	
Field Dates:	
Plan Type:	Category II

Approval

Title	Name	Signature	Date
Project leader	Andrew Barclay		
Project leader	Nicholas Decovich		2/19/14
Biometrician	Jim Jasper		2/25/14
Research Coordinator	Chris Habicht		
Region II research biologist	Mark Willette		1/31/14
Sport Fish Regional Research Supervisor	Jack Erickson		

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF APPENDICES

Appendix **Page**

PURPOSE

Coho salmon are harvested in both commercial and sport fisheries in Upper Cook Inlet (UCI), with a 10-year average of 186,655 fish being harvested annually by commercial fisheries (Shields and Dupuis 2013). Because coho salmon *Oncorhynchus kisutch* returns in northern Cook Inlet streams have been on the decline in recent years, there is a management need to estimate the harvest of these stocks in UCI fisheries. Genetic baselines are available for mixed stock analysis (MSA) of sockeye and Chinook salmon harvest samples collected from commercial, sport, and personal use fisheries in UCI; however, a genetic baseline for coho salmon in UCI has not been developed. A comprehensive coho salmon genetic baseline in Cook Inlet will allow for MSA of coho salmon harvests in UCI fisheries.

OBJECTIVES

1. Collect genetic tissue samples from at least 100 coho salmon from 10–15 spawning aggregations within Cook Inlet currently unrepresented in Alaska Department of Fish and Game (department) tissue archives. Up to 500 additional tissue samples may be collected from spawning aggregations represented in the archives to increase sample sizes and test for among-year variation.
2. Develop a genetic baseline and determine potential reporting groups for Cook Inlet coho salmon useful for estimating the stock compositions of samples collected from Upper Cook Inlet commercial and test fisheries.

BACKGROUND

Since 1992, the US Fish and Wildlife Service (USFWS) collected genetic samples from coho salmon spawning locations within Cook Inlet, with a majority being collected in Kenai and Kasilof river drainages (Table 1). In the early 2000's, the USFWS Conservation Genetics Laboratory developed a statewide baseline, which included 8 UCI coho salmon populations for 9 microsatellite loci (Olsen et al. 2003). This baseline demonstrated that genetic markers could be used to distinguish coho salmon populations in Alaska, and the possibility for distinguishing among some stocks within Cook Inlet. However, this baseline did not adequately characterize all populations that might be harvested in UCI fisheries.

Coho salmon have also been collected near or on spawning aggregations opportunistically throughout UCI by department staff since the early 1990's, with the majority collected between 2006 and 2012 (Table 1). In 2013 the state funded a 3-phase study to develop a Cook Inlet coho salmon baseline and apply this baseline to analyze fishery mixtures. The first phase involved an initial analysis using existing samples and genetic markers to determine whether the genetic diversity among Cook Inlet coho salmon populations would allow for accurate MSA estimates and was completed in spring of 2013 (Tables 1 and 2). Statistical analysis of these data indicated that sufficient variation exists in Cook Inlet coho salmon stocks for genetic stock identification.

The second phase involves collecting samples of coho salmon from additional spawning locations in Cook Inlet, analyzing their tissues for genetic markers and building and testing the baseline for MSA of UCI coho salmon. This phase began in summer of 2013 and continues through 2014. During the 2013 field season, samples of coho salmon were collected from spawning locations in Cook Inlet by several projects (Table 1): Susitna Hydroelectric Project

(499 individuals), 6 Division of Sport Fish weirs (608 individuals), Grant Creek Hydroelectric Project (100 individuals), and this project (1,899 individuals).

This operational plan includes the remainder of phase 2: sampling in the summer and fall of 2014 and laboratory and statistical analyses. Additional locations and previously-sampled locations where sample size targets were not achieved earlier will be sampled in 2014. A subset of samples will be screened for SNP loci, as in Phase 1. Statistical analyses of these samples will identify SNP loci to include in a baseline appropriate for analyze of UCI coho salmon catches, identify reporting groups, and test the baseline for the MSA performance.

The third phase of this project will occur after the baseline is built and tested. This phase will analyze approximately 5,200 fish per year from the Central District Drift Gillnet, General Subdistrict Set Gillnet, and Eastern Subdistrict Set Gillnet fisheries. In addition, approximately 800 fish per year will be analyzed from the Northern and Southern Offshore Test fisheries. The collection of fishery samples for this phase began in 2013 and will continue through 2015.

METHODS

BASELINE SAMPLING

Sample collection sizes

The ideal sample size for baseline collections to investigate population structure using markers with two alleles (i.e. single nucleotide polymorphisms [SNPs]) is 100 fish per population. This is also good target sample size for baseline populations used in MSA (Waples 1990). However, sample sizes as small as 50 fish per population may be adequate to conduct coarse-scale population structure analyses and MSA using SNPs (Seeb 2000). A population is defined as a spawning aggregate of a randomly mating group of fish that are largely reproductively isolated from other spawning aggregates.

Sampling locations

In 2013, sampling crews from several agencies and organizations collected tissue samples from 2,769 coho salmon representing putative spawning aggregates (locations) from throughout Cook Inlet (Table 1); from Kamishak and Kachemak bays in the south to the upper reaches of the Susitna River drainage in the north. Sampling crews will continue to collect coho salmon genetic tissue samples from a target of 10 to 15 locations throughout Cook Inlet between late August and mid-October, 2014 (See possible target collection locations in Table 1).

Tissue sampling

Coho salmon will be captured using either hook-and-line or seine, gill, or dip nets depending on the size of the stream and location of fish. Upon capture, a single axillary process will be clipped from each coho salmon and placed in a bottle of denatured ethyl alcohol for preservation (Appendix A1). Fish will be held in the water as much as possible while hooks are removed and samples are collected, and released immediately after the sample has been placed in the bottle. If necessary, crews will hold a fish in the water to make sure it can swim before release. Depending collection needs and project resource and Sport Fish Division staff availability, area personnel may assist in sampling efforts. Project resources will likely be available to cover Sport Fish Division costs associated with sampling and may be available for personnel costs. Resource allocation will be determined inseason on a case by case basis.

LABORATORY ANALYSIS

DNA will be extracted from axillary processes using DNeasy® 96 Tissue Kits by QIAGEN® (Valencia, CA). Samples will be analyzed for up to 96 single nucleotide polymorphism (SNP) markers that were identified as variable in Phase 1(Table 2).

The DNA samples will be analyzed using Fluidigm® 96.96 Dynamic Arrays (http://www.fluidigm.com). The Fluidigm® 96.96 Dynamic Array contains a matrix of integrated channels and valves housed in an input frame. On one side of the frame, there are 96 inlets to accept DNA extracts from individual fish and on the other are 96 inlets to accept the assay cocktails for each SNP marker. Once in the wells, the components are pressurized into the chip using the IFC Controller HX (Fluidigm®). The 96 samples and 96 assays are then systematically combined into 9,216 parallel reactions. Each reaction is a mixture of 4 microliters (µl) of assay mix (1x DA Assay Loading Buffer [Fluidigm®], 10x TaqMan® SNP Genotyping Assay [Applied Biosystems], and 2.5x ROX [Invitrogen]), and 5 µl of sample mix (1x TaqMan® Universal Buffer [Applied Biosystems], 0.05x AmpliTaq® Gold DNA Polymerase [Applied Biosystems], 1x GT Sample Loading Reagent [Fluidigm®], and 60-400ng/ul DNA) combined in a 6.7 nanoliter (nL) chamber. Thermal cycling is performed on an Eppendorf IFC Thermal Cycler as follows: an initial "hot mix" for 30 minutes at 70°C, then denaturation for 10 minutes at 96°C followed by 40 cycles of 96°C for 15 seconds and 60°C for 1 minute. The Dynamic Arrays are read on a BioMark™ Real-Time PCR System (Fluidigm®) after amplification and scored using Fluidigm® SNP Genotyping Analysis software.

For some SNP markers, genotyping will be performed in 384-well reaction plates. Each reaction is conducted in a 5 µL volume consisting of 5–40 ng of template DNA, 1x TaqMan® Universal PCR Master Mix (Applied Biosystems), and 1x TaqMan® SNP Genotyping Assay (Applied Biosystems). Thermal cycling is performed with a Dual 384-well GeneAmp PCR System 9700 (Applied Biosystems) as follows: an initial denaturation of 10 minutes at 95°C, followed by 50 cycles of 92°C for 1 second, and annealing/extension temperature for 1.0 or 1.5 minutes. The plates are scanned on an Applied Biosystems Prism 7900HT Sequence Detection System after amplification and scored using Applied Biosystems' Sequence Detection Software (SDS) version 2.2.

Genotypes collected will be entered into the GCL Oracle database, LOKI. Quality control measures will include re-extraction of 8 percent of each collection and re-analysis for all markers to ensure that genotypes are reproducible and to identify laboratory errors and rates of inconsistencies. Genotypes are assigned to individuals using a double-scoring system.

STATISTICAL ANALYSIS

Data retrieval and quality control

Genotypes will be retrieved from LOKI and imported into *R* (R Development Core Team 2011) with the *RODBC* package (Ripley 2010). Subsequent analyses will be performed in *R*, unless otherwise noted.

Prior to statistical analysis, 4 analyses will be performed to confirm the quality of the data. First, SNP markers will be identified that are invariant, or with only very low frequencies of variant alleles. These markers will be excluded from further statistical analyses.

Second, individuals will be identified that are missing substantial genotypic data, because they likely have poor-quality DNA. Individuals missing substantial genotypic data will be identified using the 80% rule (missing data at 20% or more of loci; Dann et al. 2009). These individuals will be removed from further analyses. The inclusion of individuals with poor-quality DNA might introduce genotyping errors into the baseline and reduce the accuracies of population-specific genotype frequencies.

Third, individuals with duplicate genotypes will be identified and removed from further analyses. Duplicate genotypes can occur as a result of sampling or extracting the same individual twice, and will be defined as pairs of individuals sharing the same alleles in 95% of screened loci. The individual sample with the most missing genotypic data from each duplicate pair will be removed from further analyses. If both samples have the same amount of genotypic data, the first sample will be removed from further analyses.

The final quality control analysis will identify individuals from juvenile collections that appear to be siblings (full or half siblings). Inclusion of siblings provides inappropriately precise estimates of allele frequencies. We will use the program ML-Relate (Kalinowski et al. 2006) to detect siblings and may exclude from the baseline all but one individual from every set of siblings identified, if deemed necessary.

Hardy-Weinberg expectations

For each locus within each collection, tests for conformance to Hardy-Weinberg expectations (HWE) will be performed using Monte Carlo simulation with 10,000 iterations in the *Adegenet* package (Jombart 2008). Probabilities will be combined for each collection across loci and for each locus across collections using Fisher's method (Sokal and Rohlf 1995), and collections and loci that violated HWE after correcting for multiple tests with Bonferroni's method (α = 0.05) will be excluded from subsequent analyses.

Temporal variation

Temporal variation of allele frequencies will be examined with a hierarchical, three-level analysis of variance (ANOVA). Temporal samples will be treated as sub-populations based on the method described in Weir (1996). This method will allow for the quantification of the sources of total allelic variation and permit the calculation of the among-years component of variance and the assessment of its magnitude relative to the among-population component of variance. This analysis will be conducted using the software package *GDA* (Lewis and Zaykin 2001).

Pooling collections into populations

When appropriate, collections will be pooled to obtain better estimates of allele frequencies following a step-wise protocol. First, collections from the same geographic location, sampled at similar calendar dates but in different years, will be pooled, as suggested by Waples (1990). Then differences in allele frequencies between pairs of geographically proximate collections that were collected at similar calendar dates and that might represent the same population will be tested. Collections within the same tributary (or river for mainstem spawners) will be defined as being "geographically proximate". Fisher's exact test (Sokal and Rohlf 1995) of allele frequency homogeneity will be used, and decisions will be based on a summary across loci using Fisher's method. Collections will be pooled when tests indicate no difference between collections ($P >$ 0.01). When all individual collections within a pooled collection are geographically proximate

to other collections within the same tributary, the same protocol will be followed until significant differences are found between the pairs of collections being tested. After this pooling protocol, these final collections will be considered populations. Finally, populations will be tested for conformance to HWE following the same protocol described above to ensure that pooling was appropriate, and that tests for linkage disequilibrium will not result in falsely positive results due to departure from HWE. Populations that depart from HWE will either be split into component populations or excluded from further analysis.

Linkage disequilibrium

Linkage disequilibrium between each pair of nuclear markers will be tested in each population to ensure that subsequent analyses are based on independent markers. The program *Genepop* version 4.0.11 (Rousset 2008) will be used with 100 batches of 5,000 iterations for these tests. The frequency of significant linkage disequilibrium between pairs of SNPs ($P < 0.05$) will then be summarized. Pairs will be considered linked if they exhibited significant linkage in more than half of all populations.

Hierarchical log-likelihood ratio tests

Genetic diversity will be examined with a hierarchical log-likelihood ratio (G) analysis with the package *hierfstat* (Goudet 2006).

Visualization of genetic distances

Two approaches will be used to visualize genetic distances among collections. Both approaches are based on pairwise F_{ST} estimates from the final set of independent markers with the package *hierfstat*. The first approach is to construct 1,000 bootstrapped neighbor-joining (NJ) trees by resampling loci with replacement to assess the stability of tree nodes. The consensus tree will be plotted with the *APE* package (Paradis et al. 2004). While these trees provide insight into the variability of the genetic structure of collections, pairwise distances visualized in three dimensions are more intuitive. In a second approach, pairwise F_{ST} will be plotted in a multidimensional scaling (MDS) plot using the package *rgl* (Adler and Murdoch 2010).

Assessing reporting groups for MSA

A comprehensive analysis will be conducted when SNP data are available from baseline collections sampled through 2014. We will use three methods to assess the utility of reporting groups for MSA once these data are available: 100% proof tests, the ONCOR leave-one-out method (Anderson et al. 2008), and inriver mixture samples. For the 100% proof tests, we will sample without replacement 400 individuals from each reporting group, where samples are drawn from each population within a reporting group in proportion to their population sample size in the baseline. We will estimate the stock compositions of these mixed composition proof tests and compare these estimates to the true proportions. To account for sampling error, we replicate this procedure 10 times in a manner similar to Habicht and Dann (2012).

For the leave-one-out method, we will use ONCOR, an MS Windows-based program available at http://www.montana.edu/kalinowski, to implement the simulations. This program handles only diploid markers, so we will exclude linked and mtDNA loci from the analysis. The output from this analysis produces stock proportion point estimates for each population by reporting group.

For the inriver mixture test, we will construct a mixture of 200 randomly selected coho salmon samples from 2 Susitna River fish wheel collections and 1 collection from the Deshka River weir

(Table 3). This mixture, composed entirely of Susitna River fish, will act as a 100% mixture test for the reporting group that contains the Susitna River drainage. We will compare the stock composition of this mixture to the true proportion; 100% Susitna River fish in this case.

These three analyses will determine whether the population structure is adequate for MSA to produce useful results. Generally, correct assignments of 90% to reporting groups are considered adequate for MSA.

Estimating stock composition of proof test and inriver samples

The stock compositions of the 100% proof test and inriver samples will be estimated using a Bayesian approach to genetic MSA, the Pella-Masuda Model (BAYES; Pella and Masuda 2001). The Bayesian method of MSA estimates the proportion of stocks in each mixed-stock sample using 4 pieces of information: 1) a baseline of allele frequencies for each population, 2) the grouping of populations into the reporting groups desired for MSA, 3) prior information about the stock proportions of the fishery, and 4) the genotypes of fish sampled from the fishery. We will use a flat prior for these analyses.

We will run 5 independent Markov Chain Monte Carlo (MCMC) chains of 40,000 iterations with different starting values and discard the first 20,000 iterations to remove the influences of the initial start values. We will define the starting values for the first chain such that the first 1/5 of the baseline populations sum to 0.9 and the remaining populations sum to 0.1. Each chain will have a different combination of 1/5 of baseline populations summing to 0.9. We will combine the second halves of these chains to form the posterior distribution and tabulate mean estimates, 90% credibility intervals, the probability of an estimate being equal to zero, and standard deviations from a total of 100,000 iterations. For each tabulated measure, summary statistics will be based upon the raw posterior, which will be calculated to 6 significant digits.

We will also assess the within- and among-chain convergence of these estimates using the Raftery-Lewis (within-chain) and Gelman-Rubin (among-chain) diagnostics, respectively. These values measure the convergence of each chain to stable estimates (Raftery and Lewis 1996), as well as measure the variation of estimates within a chain to the total variation among chains (Gelman and Rubin 1992), respectively. If the Gelman-Rubin diagnostic for any stock group estimate is greater than 1.2 we will reanalyze the mixture with 80,000-iteration chains following the same protocol. If the Gelman-Rubin diagnostic for any stock group estimate is greater than 1.2 after this reanalysis, we will analyze the mixture with the program HWLER (Pella and Masuda 2006). HWLER is similar to BAYES in that it estimates stock compositions based upon a Bayesian model, but differs in that it incorporates information about the effect of assigning mixture individuals to baseline populations with respect to the Hardy-Weinberg and linkage equilibria conditions observed in the baseline populations. In doing so, it allows for the identification of extra-baseline individuals that contravene equilibria conditions, but contribute to the mixture in question. We will incorporate this information into the definition of the posterior for those mixtures that failed to converge after reanalysis with 80,000-iteration chains in BAYES.

SCHEDULE AND DELIVERABLES

Preparations for the sampling season will begin in April 2014, and sampling efforts will begin approximately August 15 and end approximately September 30. Sampling preparations will include:

1) April – contacting flight services, conferring with regional staff on possible sampling locations.
2) May – purchase of sampling supplies, securing contracts with flight services.
3) June – hiring field personnel, finalizing list of potential locations.

Raw field data will be entered and error checked by October 31, 2014. Laboratory and data analyses will begin in November 2014. An ADF&G Fishery Manuscript Report will be published at the culmination of the project and will include results from the baseline analysis through 2014, due in the spring of 2015.

RESPONSIBILITIES

Andrew Barclay, Fishery Biologist III

Duties: Coordinate laboratory analysis and perform statistical analyses. Lead writing operational plans and final report. Track budgets.

Nicholas DeCovich, Fishery Biologist III

Duties: Coordinate field sampling. Contribute to writing operational plans and final report. Track budgets.

Chris Habicht, Fisheries Geneticist III

Duties: Review operational plans and reports and prioritize resources among laboratory projects to meet deadlines.

Jim Jasper, Biometrician III

Duties: Biometric support. Assist in report writing. Also reviews operational plan and final report.

Jack Erickson, Fishery Biologist IV

Duties: Coordinate collaborative sampling efforts with area Sport Fish Division staff.

REFERENCES CITED

Adler, D., and D. Murdoch. 2010. rgl: 3D visualization device system (OpenGL). R package version 0.91. http://CRAN.R-project.org/package=rgl.

Anderson, E.C., Waples, R.S., Kalinowski, S.T., 2008. An improved method for estimating the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65:1475–1486.

Buchholz WG, S. J. Miller, W. J. Spearman. 2001. Isolation and characterization of chum salmon microsatellite loci and use across species. Animal Genetics. 32:162–165.

Dann, T. H., C. Habicht, J. R. Jasper, H. A. Hoyt, A. W. Barclay, W. D. Templin, T. T. Baker, F. W. West, and L. F. Fair. 2009. Genetic stock composition of the commercial harvest of sockeye in Bristol Bay, Alaska, 2006–2008. Alaska Department of Fish and Game, Fishery Manuscript Series No. 09-06, Anchorage.

Campbell, N. R. and S. R. Narum. 2011. Development of 54 novel single-nucleotide polymorphism (SNP) assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. Molecular Ecology Resources. 11(Suppl. 1):20-30.

Cairney M., J. B. Taggart, and B. Høyheim. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar L.*) and cross-species amplification in other salmonids. Molecular Ecology 9:2155–2234.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–511.

Goudet, J. 2006. hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-4. http://www.r-project.org, http://www.unil.ch/popgen/softwares/hierfstat.htm

Greig, C. and M. A. Banks. 1999. Five multiplexed microsatellite loci for rapid response run identification of California's endangered winter Chinook salmon. Animal Genetics 30:316–324.

Habicht, C., and T. H. Dann. 2012. Western Alaska Salmon Stock Identification Program Technical Document 27: Sockeye salmon reporting group evaluations using simulated fishery mixtures. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J12-27, Anchorage. http://www.adfg.alaska.gov/FedAidPDFs/RIR.5J.2012.27.pdf (Accessed December 13, 2012).

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405. doi: 10.1093/bioinformatics/btn129.

Kalinowski, S. T., A. P. Wagner, and M. L. Taper. 2006. M-relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes, 6(2), 576-579.

Lewis, P. O. and D. Zaykin. 2001. Genetic data analysis: computer program for the analysis of allelic data. Version 1.0. URL http://lewis.eeb.uconn.edu/lewishome/software.html.

Nelson R. J. and T. D. Beacham. 1999. Isolation and cross species amplification of microsatellite loci useful for study of Pacific salmon. Animal Genetics 30:228–229.

Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7:1087-1089.

Olsen, J. B., S. J. Miller, W. J. Spearman, and J. K. Wenburg. 2003. Patterns of intra- and inter-population genetic diversity in Alaskan coho salmon: Implications for conservation. Conservation Genetics 4:557–569.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

Pella, J., and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin 99(1):151–167.

Pella J., and M. Masuda. 2006. The Gibbs and split-merge sampler for population mixture analysis from genetic data with incomplete baselines. Canadian Journal of Fisheries and Aquatic Sciences 63:576–596.

REFERENCES CITED (Continued)

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Raftery, A. E., and S. M. Lewis. 1996. Implementing MCMC. Pages 115–130 [In] W. R. Gilks, S. Richardson, and D.J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall, Inc., London

Ripley, B. 2010. RODBC: ODBC Database Access. R package version 1.3-2. http://CRAN.Rproject.org/package=RODBC.

Rousset, F. 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8(1):103–106.

Scribner K. T., J. R. Gust, R. L. Fields. 1996. Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. Canadian Journal of Fisheries and Aquatic Sciences 53:833–841.

Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R Z. Davis, L. K. Brannian, and J. E. Seeb. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the Exxon Valdez oil spill. Transactions of the American Fisheries Society 129:1223–1249.

Shields, P. and A. Dupuis. 2013. Upper Cook Inlet commercial fisheries annual management report, 2012. Alaska Department of Fish and Game, Fishery Management Report No. 13-21, Anchorage.

Small, M. P., T. D. Beacham, R. E. Withler, and R. J. Nelson. 1998. Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia. Molecular Ecology 7:141–155.

Smith C. T., B. F. Koop, R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1614–1616.

Smith C. T., L. Park, D. Vandoornik, L. W. Seeb, and J. E. Seeb. 2006. Characterization of 19 single nucleotide polymorphism markers for coho salmon. Molecular Ecology 6:715-720.

Sokal, R. R. and F. J. Rohlf. 1995. *Biometry*. 3rd Edition. Freeman, San Francisco, CA.Thomas, A., O'Hara, B., Ligges, U., and Sturtz, S. 2006. Making BUGS Open. R News 6 (1): 12- 17.

Waples, R. S. 1990. Temporal changes of allele frequency in Pacific salmon: implications of mixed-stock fishery analysis. Canadian Journal of Fisheries and Aquatic Sciences 47:968–976.

Waples R.S. 1991. Genetic methods for estimating the effective size of cetacean populations. In: (ed. Hoezel AR) Report of the International Whaling Commission, pp. 279–300. International Whaling Commission, UK.

Weir, B. 1996. Genetic Data Analysis (second edition). Sinauer Associates, Inc, Sunderland, MA.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358–1370.

TABLES

Table 1.–Cook Inlet Coho baseline and mixture sampling locations, number of archived samples (N), number of samples needed to reach a total of 100 samples for a location (Need), the number of samples analyzed in Phase 1, and the source of the collection. Of the 5,955 samples in the "Need" column, 1,500-2,000 samples are anticipated to be collected in 2014.

Area/ Drainage	Location	Year Collected	N	Need	Phase 1	Source
West Side						
	Douglas River	2013	106	--	--	This project
	Douglas Reef River	2013	113	--	--	This project
	Kamishak River	2013	110	--	--	This project
	Little Kamishak River	2013	96	4	--	This project
	McNeil River	2013	41	59	--	This project
	Sunday Creek	2012	7	93	--	This project
	Brown's Peak Creek	2013	9	91	--	This project
	Fitz Creek	2013	3	97	--	This project
	Tuxedni River	2012	86	14	81	ADF&G Archives
	Crescent Lake - Late	1998	99	--	95	USFWS
	Crescent River	2012	1	--	--	ADF&G Archives
		2013	131		--	This project
	Harriet Creek	2012	1	99	--	ADF&G Archives
	Packers Creek	2013	4	96	--	This project
	Little Jack	2013	104	--	--	This project
	Montana Bill Creek	2012	101	--	95	ADF&G Archives
	Big River	2009	19	81	--	ADF&G Archives
	Kustatan River	2013	119	--	--	This project
	Farros Lake Outlet Creek	2013	17	83	--	This project
	Nikolai Creek	--	0	100	--	--
	Chuitna River	1992	54	46	--	USFWS
	Wilson Creek	2010	223	--	94	ADF&G Archives
	Middle Creek	2008	40	60	--	ADF&G Archives
	Lone Creek	2008	70	30	--	ADF&G Archives
	Coal Creek	2013	41	59	--	This project
	Theodore River weir	2012	19	21	--	Sport Fish weir
		2013	60		--	Sport Fish weir
	Lewis River weir	2013	57	43	--	Sport Fish weir
Susitna River Drainage						
	Indian River	2013	104	--	--	SuHydro
	Susitna River - Slough 11	2013	1	99	--	SuHydro
	Whiskers Creek	2013	79	21	--	SuHydro
	Honolulu Creek	2013	4	96	--	SuHydro
	Spink Creek	2008	38	62	--	ADF&G Archives
	Troublesome Creek	2013	92	8	--	SuHydro
	Bunco Creek	2013	9	91	--	SuHydro

-continued-

Table 1.–Page 2 of 4.

Area/ Drainage	Location	Year Collected	N	Need	Phase 1	Source
Susitna River Drainage						
	Swan Lake	2009	20	80	--	ADF&G Archives
	Iron Creek	2013	28	72	--	SuHydro
	Sheep River	2013	115	--	--	SuHydro
	Larson Creek	2011	84	16	84	ADF&G Archives
	Chunilna Creek (Clear Creek)	2013	66	34	--	SuHydro
	Fish Creek	2013	1	99	--	SuHydro
	Answer Creek	2013	7	93	--	This project
	Question Creek	2013	77	23	--	This project
	Montana Creek weir	2013	200	--	--	Sport Fish weir
	Sheep Creek	--	0	100	--	--
	Kashwitna River	--	0	100	--	--
	Little Willow Creek	--	0	100	--	--
	Willow Creek	--	0	100	--	--
	Moose Creek (Deshka River)	--	0	100	--	--
	Kroto Creek	--	0	100	--	--
Yentna River Drainage						
	West Fork Yentna River	--	0	100	--	--
	Cache Creek	--	0	100	--	--
	Martin Creek	2013	35	65	--	This project
	Sunflower Creek	--	0	100	--	--
	Kichatna River	--	0	100	--	--
	Red Creek	--	0	100	--	--
	Hewitt Creek	--	0	100	--	--
	Happy River	--	0	100	--	--
	Canyon Creek	2008	20	25	--	This project
		2013	55		--	This project
	Talachulitna River	2013	74	26	--	This project
	Shell Creek	--	0	100	--	--
Knik Arm						
	Little Susitna River weir	2013	97	3	--	Sport Fish weir
	Fish Creek weir	2009	203	--	93	Sport Fish weir
		2013	94		--	Sport Fish weir
	Wasilla Creek	2013	9	91	--	This project
	Cottonwood Creek	--	0	100	--	--
	Rabbit Slough	2011	95	5	95	ADF&G Archives
	Granite Creek	--	0	100	--	--
	Moose Creek	--	0	100	--	--

-continued-

Table 1.–Page 3 of 4.

Area/ Drainage	Location	Year Collected	N	Need	Phase 1	Source
Knik Arm						
	Eska Creek	2013	61	39	--	This project
	Matanuska River mainstem	2008	135	--	--	USFWS
		2009	194		94	USFWS
	Jim Lake	2011	7	93	--	ADF&G Archives
	Jim Creek	2009	68	32	68	ADF&G Archives
	Eagle River	--	0	100	--	--
	Sixmile Creek	2009	46	54	45	ADF&G Archives
	Chester Creek	2011	54	46	53	ADF&G Archives
	Ship Creek	1991	11	89	--	ADF&G Archives
		2012	400	--	93	ADF&G Archives
Turnagain Arm						
	Campbell Creek	1995	5		--	ADF&G Archives
		2009	125	--	95	ADF&G Archives
		2010	9		--	ADF&G Archives
	Rabbit Creek	2011	54	46	53	ADF&G Archives
	Twentymile River	--	0	100	--	--
	Williwaw Creek	2013	22	78	--	This project
	Portage Creek	2013	5	95	--	This project
	Explorer Pond	2013	94	6	--	This project
	Ingram Creek	2013	7	93	--	This project
	Sixmile Creek	--	0	100	--	--
	Resurrection Creek	2010	96	4	93	ADF&G Archives
	Mystery Creek	2010	22	78	20	ADF&G Archives
	Chickaloon River	2010	82	18	80	ADF&G Archives
Northwestern Kenai Peninsula						
	Sucker Creek (Swanson River trib)	1997	94	6	91	USFWS
	Swanson River mainstem	--	0	100	--	--
	Gruska Creek (Swanson River trib)	2013	53	47	--	This project
	Bishop Creek	--	0	100	--	--
Kenai River Drainage						
	Grant Creek weir	2013	100	--	--	Grant Hydro
	Snow River - South Fork	1998	73	--	71	USFWS
		2002	50		24	USFWS
	Trail Creek	2006	134	--	--	USFWS
	Summit Creek/Quartz Creek	1998	75	25	--	USFWS
	Summit Creek	2002	50	50	--	USFWS
	Moose Creek - Kenai River	1993	150	--	--	ADF&G Archives

-continued-

Table 1.–Page 4 of 4.

Area/ Drainage	Location	Year Collected	N	Need	Phase 1	Source
Kenai River Drainage						
	below Kenai Lake (mainstem)	1999	56	--	--	USFWS
		2002	57		--	USFWS
	Russian River	2002	31	--	--	USFWS
		2013	101		--	This project
	Skilak Lake - Upper	1999	60	40	60	USFWS
	Skilak River	2003	100	--	--	USFWS
	Skilak Lake - Lower	1999	20	80	18	USFWS
	below Skilak Lake (mainstem)	1999	20		18	USFWS
		1999	60	--	60	USFWS
		2002	50		--	USFWS
	Killey River	2000	68	--	67	USFWS
		2002	49		25	USFWS
	East Fork Moose River	2000	11	--	--	USFWS
		2002	100		--	USFWS
	Moose River weir	1998	35	65	--	USFWS
	Funny River	2006	150	--	--	USFWS
	Soldotna Creek	2013	8	92	--	This project
	Slikok Creek	2008	67	33	--	USFWS
	Beaver Creek	2013	12	88	--	This project
Kasilof River Drainage						--
	Glacier Creek	2009	68	32	--	USFWS
	Indian Creek	2009	55	45	--	USFWS
	Shantatalik Creek	2009	41	59	--	USFWS
	Nikolai Creek	2009	92	8	88	USFWS
	Kasilof Mainstem	2009	100	--	--	USFWS
	Crooked Creek	--	0	100	--	--
	Coal Creek	--	0	100	--	--
Southern Kenai Peninsula						
	Ninilchik River	2013	108	--	--	This project
	Deep Creek	2013	101	--	--	This project
	Anchor River weir	2006	164	--	55	Sport Fish weir
		2009	40		40	Sport Fish weir
	Stariski Creek	2013	59	41	--	This project
	Fox River	2013	100	--	--	This project
	English Bay River	2013	12	88	--	This project

Table 2.–Marker type and source of coho salmon genetic markers used in Phase 1 of this study.

Marker Type[1]	Source	Marker Name[2]	Marker Type[1]	Source	Marker Name[2]
1	A	*Ogo2*	2	I	*Oki106419-292*
1	B	*Oke2*	2	I	*Oki106479-278*
1	B	*Oke3*	2	I	*Oki107336-45*
1	B	*Oke4*	2	I	*Oki107607-213*
1	C	*Oki11*	2	I	*Oki107974-46*
1	C	*Oki3*	2	I	*Oki108505-331*
1	D	*Oneμ3*	2	I	*Oki109243-480*
1	E	*Ots101*	2	I	*Oki109651-152*
1	F	*OTS105*	2	I	*Oki109874-122*
1	G	*Ots-2M*	2	I	*Oki109894-418*
1	H	*Ssa407UOS*	2	I	*Oki110064-418*
2	I	*Oki100771-83*	2	I	*Oki110078-191*
2	I	*Oki100974-293*	2	I	*Oki110689-43*
2	I	*Oki101119-1006*	2	I	*Oki111681-407*
2	I	*Oki101419-103*	2	I	*Oki113457-324*
2	I	*Oki101554-359*	2	I	*Oki114315-360*
2	I	*Oki101770-525*	2	I	*Oki114448-101*
2	I	*Oki102213-604*	2	I	*Oki114587-309*
2	I	*Oki102414-499*	2	I	*Oki116362-411*
2	I	*Oki102457-67*	2	I	*Oki116865-244*
2	I	*Oki102801-511*	2	I	*Oki117043-374*
2	I	*Oki102867-667*	2	I	*Oki117144-64*
2	I	*Oki103271-161*	2	I	*Oki117286-291*
2	I	*Oki103577-70*	2	I	*Oki117742-259*
2	I	*Oki103713-182*	2	I	*Oki117815-369*
2	I	*Oki104515-99*	2	I	*Oki118152-314*
2	I	*Oki104519-45*	2	I	*Oki118175-264*
2	I	*Oki104569-261*	2	I	*Oki118654-330*
2	I	*Oki105105-245*	2	I	*Oki94903-192*
2	I	*Oki105115-49*	2	I	*Oki95318-100*
2	I	*Oki105132-169*	2	I	*Oki96127-66*
2	I	*Oki105235-460*	2	I	*Oki96158-278*
2	I	*Oki105385-521*	2	I	*Oki96376-63*
2	I	*Oki105407-161*	2	I	*Oki97954-228*
2	I	*Oki105897-298*	2	J	*Oki_Cr-209*
2	I	*Oki106172-60*	2	J	*Oki_Cr-296*
2	I	*Oki106313-353*	2	K	*Oki_Car-353*

-continued-

Table 2.–Page 2 of 2.

Marker Type[1]	Source	Marker Name[2]	Marker Type[1]	Source	Marker Name[2]
2	J	*Oki_E2-87*	2	L	*Oki_hsc71p-313*
2	J	*Oki_GPDH-146*	2	L	*Oki_hsf1b-85*
2	J	*Oki_GPDH-188*	2	K	*Oki_il1rac-169*
2	J	*Oki_GnRH-151*	2	J	*Oki_ins-167*
2	J	*Oki_HGFA-311*	2	J	*Oki_ins-323*
2	J	*Oki_IGF-I.1-163*	2	L	*Oki_itpa-85*
2	J	*Oki_LWSop-554*	2	L	*Oki_metA-220*
2	J	*Oki_il-1racp-176*	2	L	*Oki_nips-159*
2	J	*Oki_SClkF2R2-120*	2	L	*Oki_p53-20*
2	L	*Oki_SECC22-67*	2	L	*Oki_parp3-19*
2	J	*Oki_SWS1op-38*	2	L	*Oki_pigh-33*
2	K	*Oki_TniUPP-230*	2	L	*Oki_pop5-265*
2	K	*Oki_U202-136*	2	L	*Oki_rpo2j-235*
2	K	*Oki_U202-258*	2	J	*Oki_serpin-130*
2	K	*Oki_U216-151*	2	J	*Oki_serpin-328*
2	J	*Oki_arf-115*	2	L	*Oki_spf30-119*
2	L	*Oki_arp-105*	2	L	*Oki_srp09-107*
2	L	*Oki_aspAT-273*	2	L	*Oki_sys1-141*
2	L	*Oki_bcAKal-274*	2	L	*Oki_taf12-40*
2	L	*Oki_carban-140*	2	L	*Oki_txnip-35*
2	J	*Oki_eif4ebp2-58*	2	J	*Oki_u6-257*
2	L	*Oki_gdh-189*	2	L	*Oki_vatf-363*
2	L	*Oki_gh-183*			

[1] Marker type: 1) microsatellite; 2) single nucleotide polymorphism.

2 Marker source: A) Olsen et al. (1998); B) Buchholz et al. (2001); C) Smith et al. (1998); D) Scribner et al. (1996); E) Small et al. (1998); F) Nelson and Beacham (1999); G) Greig and Banks (1999); H) Cairney et al. (2000); I) Southwest Fisheries Science Center (Unpublished); J) Smith et al. (2006); K) University of Washington (unpublished); L) Campbell and Narum (2011).

Table 3.–Available coho salmon mixture collections for inriver test mixtures including, sampling location, year collected, sample size (N), and collection source.

Location	Year Collected	N	Source
Deshka River weir	2013	100	Sport Fish Division weir
Susitna Camp Fish Wheel (West)	2013	296	Susitna Hydroelectric Project
Susitna Camp Fish Wheel (East)	2013	296	Susitna Hydroelectric Project

APPENDIX A
GENETIC SAMPLING INSTRUCTIONS

Non-lethal Bulk Sampling Finfish Tissues for DNA Analysis

ADF&G Gene Conservation Lab, Anchorage

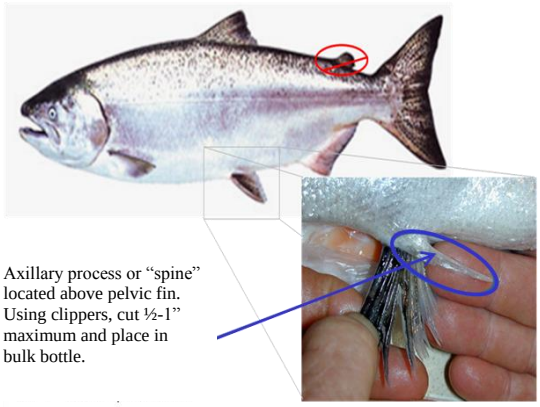
I. General Information

We use axillary process samples from individual fish to determine the genetic characteristics and profile of a particular run or stock of fish. This is a non-lethal method of collecting tissue samples from adult fish for genetic analysis. The most important thing to remember in collecting samples is that **only quality tissue samples give quality results**. If sampling from carcasses: tissues need to be as "fresh" and as cold as possible and recently moribund, do not sample from fungal fins.

II. Sampling Method

Preservative used: Isopropanol/Methanol/Ethanol (EtOH) preserves tissues for later DNA extraction. Avoid extended contact with skin.

Sampling instructions are written for (N=100 fish/125ml) bulk bottle. <u>Steps for collecting axillary process tissues:</u>



Axillary process or "spine" located above pelvic fin. Using clippers, cut ½-1" maximum and place in bulk bottle.

- Wipe dry the axillary process "spine" prior to sampling to avoid getting excess water or fish slime into the 125ml bottle (see diagram).
- Clip off the axillary "spine" using dog nail clippers or scissors to get roughly a **½ - 1" inch maximum** piece and/or about the size of a small fingernail.
- Place each tissue piece into bulk bottle **(place only one piece of axillary from each fish)**.
- Repeat: **up to 100 fish /125ml bulk bottle (**into same bottle). If you don't reach this number of fish per location, that's ok. Maximum storage capacity 125ml bulk for proper preservation of axillary tissue is (N=100).
- Record on **each label:** Location, sampling date (mm/dd/yyyy), sampler's name(s), total number of fish sampled, latitude/longitude, and field notes (if any). Use pencil. This insures correct data with each collection bottle.
- If collection occurs over 4~5 day period, "refresh" EtOH at end of the collection.
- After the collection is complete and 24 hours have passed, "refresh" the axillary tissues as follows: carefully pour off ¾ EtOH and then pour fresh EtOH into sample bottle containing axillary clips. Cap and invert bottle twice mixing EtOH and tissue.
- Freezing not required, store sample bottle in upright cool location for good tissue quality.

SILLY: _____
Location: _____
Sample Date(s):___/___/___
Sampler's name:_____
Total # fish sampled:_____
Latitude:_____
Longitude:_____
Species:_____
Comments:_____
ADF&G:Preserved in EtOH

<u>Supplies included in sampling kit</u>:

1. Clipper- used to cut a portion of **one** axillary process per fish.
2. Sample target: 100 axillary clips/125ml bulk bottle.
3. Labels on bulk sample bottles: Location, Sample date, Sampler, Total # fish sampled and comments (if any).
4. **1:125ml** wide mouth bottle(s) for EtOH "refresh" step.
5. Sampling instructions.

Return to ADF&G Anchorage lab:	ADF&G – Genetics 333 Raspberry Road Anchorage, Alaska 99518	Lab staff: 907-267-2247 Judy Berger: 907-267-2175 **Freight code:** _____

Appendix A1.– Bulk sampling instructions for adult salmon. Fin tissue will be sampled when axillary process is not available.